



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/714,690

11/17/2003

Elizabeth V. Bagley

RSW920030237US1 (129)

5361

46320

7590

06/26/2007

CAREY, RODRIGUEZ, GREENBERG & PAUL, LLP

STEVEN M. GREENBERG

950 PENINSULA CORPORATE CIRCLE

SUITE 3020

BOCA RATON, FL 33487

EXAMINER

SHAH, PARAS D

ART UNIT

PAPER NUMBER

2626

MAIL DATE

DELIVERY MODE

06/26/2007

PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

## Office Action Summary

Application No.

10/714,690

Applicant(s)

BAGLEY ET AL.

Examiner

Paras Shah

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 17 November 2003.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-12 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-12 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 17 November 2003 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- ☒ Notice of References Cited (PTO-892)
- ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- ☒ Information Disclosure Statement(s) (PTO/SB/08)  
Paper No(s)/Mail Date 11/17/2003
- ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- ☐ Notice of Informal Patent Application
- ☐ Other: \_\_\_\_\_

### **DETAILED ACTION**

1. This communication is in response to the Application filed on 11/17/2003. Claims 1-12 are pending and have been examined.

#### ***Information Disclosure Statement***

2. The information disclosure statement (IDS) submitted on 11/17/2002 was filed in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

#### ***Specification***

3. The disclosure is objected to because of the following informalities: "sting" should be "string" on page 6, [0009], line 7.

Appropriate correction is required.

#### ***Claim Rejections - 35 USC § 102***

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. Claim 1, 3, 7, 8, and 12 are rejected under 35 U.S.C. 102(b) as being anticipated by Kiyama *et al.* (US 5,642,518)

As to claim 1, Kiyama *et al.* discloses a keyword generation system comprising:

a content parser configured to parse individual words and phrases in a selected portion of content (see col. 4, lines 58-60) (e.g. The dividing of the words from the text is equivalent to a parser);

a dictionary of words and phrases (see col. 4, line 60) (e.g. A dictionary is being used when words are divided.) specific to a particular domain associated with said content (see col. 1, lines 4-5) (e.g. It is implied in the reference that keywords associated with a domain type is extracted and is thus specific to a particular domain depending on the word detected).;

a list of keyword candidates comprising a plurality of words and phrases specific to said particular domain (see col. 4, lines 66-67) (e.g. A keyword list is generated of the detected words corresponding to specific domains (see Figure 9);

a counter for each of said words and phrases in said list (see col. 4, lines 64-65) (e.g. a frequency of occurrence is equivalent to a counter since both are directly proportional to each other. Further, once a keyword is detected the count increases (see Figure 9, element w3); and,

a keyword generation process (see Figure 2, element 70) both coupled to each of said content parser (see Figure 2, element 10), dictionary (see col. 4, line 60) , said list (see Figure 2, elements 20 and 30), and said counter (see Figure 2, element 20 and 30) and also programmed (see col. 6, lines 22-39) (e.g. The use of a processor implies the programming portion) to identify said words and phrases specific to said particular domain in said selected portion of content (see

Figure 2, element 30 and col. 5, lines 1-4 and Figure 9) (e.g. From Figure 9, it can be seen that the occurrence frequencies are identified as well as the domain as seen by the element columns, w3), to write said identified words and phrases to said list of keyword candidates (see col. 5, lines 5-7) (e.g. The extracted keywords are stored for new keywords and existing keywords are already in the dictionary (see Figure 9, element column, w1)) , to increment said counter for each of said words and phrases in said list each time said keyword generation process locates each of said words and phrases in said selected portion of content (see col. 5, lines 48-55) (e.g. In this cited portion the, word occurrence frequencies are updated depending on the word and incremented by domain and shown in Figure 9), and to select one or more of said words and phrases in said list as keywords for said content based upon said counter for each of said words and phrases in said list (see col. 11. lines 59-65 and Figure 18) (e.g. The assignment of the keywords shows the selection of the keywords based on the occurrence frequency. The latter citation shows the keywords assigned).

As to claims 3 and 8, Kiyama *et al.* discloses a keyword generation method comprising the steps of:

locating words and phrases in a selected portion of content(see col. 4, lines 58-60), said words and phrases being specific to a particular domain (see col. 1, lines 4-5) (e.g. It is implied in the reference that keywords associated with

a domain type is extracted and is thus specific to a particular domain depending on the word detected).;;

adding a single instance of each of said located words and phrases to a list of keyword candidates (see col. 5, lines 5-7) ;

for each located word and phrase which already had been added to said list of keyword candidates, incrementing a counter associated with said located word and phrase (see col. 4, lines 64-65) (e.g. a frequency of occurrence is equivalent to a counter since both are directly proportional to each other. Further, once a keyword is detected the count increases (see Figure 9, element w3); and,

selecting keywords from said list of keyword candidates based upon words and phrases in said list (see col. 11. lines 59-65 and Figure 18) (e.g. The assignment of the keywords shows the selection of the keywords based on the occurrence frequency. The latter citation shows the keywords assigned.) having a highest counter value (see Figure 17) (e.g. From the Figure, if the keyword has not been seen more than one time then it is assigned as the keyword to the specific domain. If it is has only been seen once then it is not assigned) (see Figure 18 and Figure 16, example)).

As to claims 7 and 12, Kiyama *et al.* discloses

the step of repeated performing the locating, adding and incrementing steps for selected chunks of said selected portion of content until no content

remains to be processed (see Figure 3, elements, 1, 11-16) (e.g. From the Figure, it is seen that the text is retrieved and processing is done until the text is completed (see Figure 4, sample document)).

### ***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 2, 4, and 9 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kiyama *et al.* in view of de Hita *et al.* (US 6,081,774).

As to claim 2, Kiyama *et al.* does not specifically disclose the use of a list of common words for keyword generation.

However, de Hita *et al.* does disclose the use of list of common words (see col. 3, lines 4-8) (e.g. pattern dictionary) coupled to keyword generation process (see col. 3., lines 51-61).

It would have been obvious to one of ordinary skilled in the at the time the invention was made to have modified the key word generation taught by Kiyama *et al.* with the inclusion of word removal common to keywords as taught by de Hita *et al.* The motivation to have combined the two references involves the inclusion of context dependent information related to semantic relationships (see

Art Unit: 2626

de Hit a *et al.*, col. 3, lines 3-8) in order to merge expressions that are similar (see de Hita *et al.*, col. 2, lines 9-13).

As to claims 4 and 9, Kiyama *et al.* does not specifically disclose removing from consideration every word and phrase in said list of keyword candidates that are common.

However, de Hita *et al.* does disclose removing from consideration the keywords and words common in nature (see col. 3., lines 51-61) (e.g. From the cited sections it is implied that the synonyms for the word or token is retrieved, which is one of the problems the reference tries to solve (see col. 2, lines, 11-14).

It would have been obvious to one of ordinary skilled in the at the time the invention was made to have modified the key word generation taught by Kiyama *et al.* with the inclusion of word removal common to keywords as taught by de Hita *et al.* The motivation to have combined the two references involves the inclusion of context dependent information related to semantic relationships (see de Hit a *et al.*, col. 3, lines 3-8) in order to merge expressions that are similar (see de Hita *et al.*, col. 2, lines 9-13).

8. Claims 5, 6, 10, and 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kiyama *et al.* in view of Kim. (US 2003/0208482).

As to claims 5 and 10, Kiyama *et al.* discloses



selecting a string (e.g. word) in said selected portion of content (see Figure 3, elements 1, and 11-16) (e.g. The input text is the selection and words are extracted);

adding said string to said list of keyword candidates (see col. 5, lines 5-10) (e.g. The comparison to the collection and the negligible word dictionary is made and added if the respective word is not found in the collection or negligible word dictionary) ;

However, Kiyama *et al.*, does not specifically disclose the detecting a variation in font attributes.

Kim does disclose the detecting of words based upon font attributes (see [0031]).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have modified the key word generation taught by Kiyama *et al.* with the inclusion of font detection as taught by Kim *et al.* The motivation to have combined the two references involves the distinction between important words and unimportant words comparing other words in the text [see Kim *et al.* [0031]], which benefits the keyword generation presented by Kiyama *et al.* by detecting keywords denoted by font to be important.

As to claims 6 and 11, Kiyama *et al.*, further discloses

the step of subsequently identifying said string as a word and phrase (see col. 5, lines 48-55) , which is specific to said particular domain.(see Figure 9 and

Art Unit: 2626

16) (e.g. The occurrence frequency is used to determine the domain for which the keywords extracted from a text document belongs to).

### ***Conclusion***

9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Morimoto *et al.* (US 5,297,027) is cited to disclose textual information extraction, more specifically words. Imanaka *et al.* (US 6,064,952) is cited to disclose a keyword extraction based on character strings.

Shibata *et al.* (JP 08-221448) is cited to disclose a keyword extracting based on frequency. Yoshinaga (JP 11-203293) is cited to disclose a synonym retrieval based on input word.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Paras Shah whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:30a.m.-4:00p.m. EST.


If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571)272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

P.S.

06/13/2007

  
PATRICK N. EDOUARD  
SUPERVISORY PATENT EXAMINER